International Journal of Pharmaceutical and Bio-Medical Science

ISSN(print): 2767-827X, ISSN(online): 2767-830X

Volume 06 Issue 02 February 2026

Page No : 715-723

DOI: https://doi.org/10.47191/ijpbms/v6-i2-06, Impact Factor: 8.163

Metagenomic Next-Generation Sequencing for Viral Pathogen Detection: A Comprehensive Review

Aomnea M Albdulazeez[1], Shahad.I. Saeed[2], Mohammed T. Al-noaemy[3], Hamid, Mohammed Mahmood[4], Hameed, Amna Yarub[5]

[1]Department of anesthesia /Mosul Medical Technical Institute/ Northern Technical University Mosul, IRAQ

[2,3]Optometry Techniques Department / Mosul Medical Technical Institute/ Northern Technical University Mosul, IRAQ

[4] Emergency Medicine and Immediate Aid Techniques Department/ Mosul Medical Technical Institute/ Northern Technical University Mosul, IRAQ

[5] Clinical laboratory sciences/ Mosul Medical Technical Institute/ Northern Technical University Mosul, IRAQ

ABSTRACT

Viruses, the most abundant biological entities on Earth, have remained largely uncharted, despite their significant ecological impact. Next-generation sequencing (NGS) technologies have catalyzed a revolution in viral discovery by enabling rapid retrieval of vast RNA and DNA sequences from diverse sources. Metagenomic next-generation sequencing (mNGS) surpasses traditional methods by facilitating the detection of both known and novel viruses without the need for pre-existing knowledge of genomes. It encompasses various paradigms, including shotgun sequencing and targeted amplification, which allow for comprehensive viral detection and classification. The evolution of bioinformatics tools has bolstered mNGS applications, aiding in clinical diagnostics, environmental monitoring, and epidemiological studies. mNGS has proven particularly valuable in identifying pathogens linked to complex syndromes and emerging threats, such as SARS-CoV-2. By revealing viral genomics across clinical samples and environmental matrices, mNGS enhances our understanding of viral evolution and diversity, critical to public health surveillance and biosecurity. The advent of advanced sequencing techniques further improves detection accuracy and coverage, fostering the development of bioinformatic pipelines that refine analytical processes. Metagenomic technologies promise robust applications in clinical diagnostics and environmental virology, enabling proactive responses to emerging viral threats while addressing public health challenges through comprehensive and efficient pathogen monitoring and detection.

KEYWORDS: Metagenomic, viral detection, Next-generation sequencing

ARTICLE DETAILS

Published On:
14 February 2026

Available on:
https://ijpbms.com/

1. INTRODUCTION

Viruses represent the most abundant biological entities on our planet, yet their real-world diversity remains vastly unexplored [1]. Even though they exert a considerable influence on all ecosystems, viral discovery has lagged behind that of other microbes and is, consequently, particularly relevant for a sustainable environment [2]. Next-generation sequencing (NGS) technologies allow the rapid retrieval of millions of RNA and DNA sequences from countless sources like environmental, animal, and plant specimens, fostering yet another major revolution in this area. At its conception, metagenomic sequencing (shotgun sequencing) was merely an intellectual hypothesis along the road to archivable virome catalogues [3]. The habitual low conservation of virus genome sequences among distant relatives, combined with the impossibility of procuring a unique injection-stage or an intermediate-stage viral assembly, made individual-specific-generation sequencing impractical. With shotgun metagenomics, the circumvention of direct tracing of the non-cultivatable strain appeared achievable by unwarranted and unpretendedly extensive coverage of all constituent microorganisms within a sample [4]. Virome disentangling presently appears realistic under free-assembly frameworks aided by read piston noise, which

diminishes per-generation hubbub through a spectral information dilution. The ill-finished riddle of missing habitat-specific viriomes extant meta-NGS, i.e., virome milieu records mediating viral shifts traced along the autochthonous hymenopteran trajectory with continuous elemental or genome-wise latches, similarly awaits a feasible resolution manifestation [5].

The trajectory of the "resurrection" of mNGS, supported by co-operative "omics" as well as biochemical reaction-based disciplines, presently shows promise in systematic omnivorous population studies on various organisms and the construction of four-dimensional biofactories capable of biosynthesizing macro-bioactivity accessible across the globe [6]. Based on identifiable genomic fragments procured once containment along these lines begins, the recuperation of partial or full pan-genomic information along entanglements appears plausible [7]. Furthermore, the uses of mNGS assistance in formulating hypotheses to trace the genesis of unprecedented human disorders are likewise attainable. The surging importance of viral metagenomic NGS in clinical expositions can be traced through unexplored designated pathways and scarcely examined backwater humic kinds, which hold the rank of fourth, retaining an unfielded status trivialize previous statements on critical zoonotic-associated kinds yet crucial to public health [8].

2. PRINCIPLES OF METAGENOMIC NEXT-GENERATION SEQUENCING

Metagenomic next-generation sequencing (mNGS) has emerged as a powerful tool for interrogating viral pathogens across various matrices and biological systems, enabling identification of known and novel viruses alike. NGS technologies allow massive parallel sequencing of DNA or RNA fragments, producing millions of short reads that require ulterior assembly into longer sequences, annotation, and classification [9]. Typically, a relatively low portion of the reads reflects the nucleic acid of the target organism. mNGS constitutes an attractive alternative to classical virus detection methods based on target amplification or sequencing of selected genomic areas. In viral ecology studies, mNGS has expanded enormously from single-species reads to metagenomic datasets containing hundreds or thousands of different virus and virus-like sequences. With the concomitant development of bioinformatic tools for sequence processing, annotation, and classification, viral genomics research has shifted from pure virology to viromics, virosphere exploration, and virome studies [10]. The methods presented have been classified into five main paradigms based on library-preparation strategies and sequencing instruments:

([1]): RNA-sequencing libraries containing a poly(A) tail, the 5' cap, or random oligonucleotide primers, or cDNA amplification using reverse transcriptases for the detection of RNA viruses.

([2]): Exogenous virion-remnant enrichment, virus- and host-targeted amplification procedures, and rolling-circle amplification for the detection of double-stranded DNA (dsDNA) viruses.

([3]): DNA circularization, genome-fragmenting chemicals, and isothermal amplification techniques to target single-stranded DNA (ssDNA) viruses.

([4]): A variety of enrichment and amplification methods adapted to study viruses hosted by specific organisms.

([5]): Unbiased detection of ribosomal RNA–depleted RNA viruses, viruses associated with raw or treated wastewater, and viruses present in outdoor and indoor air.

The mNGS methodology has gained in coverage, precision, and accuracy with the emergence of long-read sequencing technologies, which also recover complete viral genomes, thereby supporting assembly, annotation, and phylogenetic studies [11].

More than [200] NGS systems worldwide are currently operational, whose DNA- and RNA-sequencing platforms support the establishment of mNGS in health and veterinary research. However, constant assessment of transparent methodologies for obtaining relevant genomes remains essential [12], [13]. Already available open-access dedicated databases, software tools, bioinformatic resources and in-depth analyses are now provided for guidance in refining surveillance activities or exploring unexpected or unexplained findings [2]. Such tools additionally enable selection of suitable sequencing techniques and devices, consequently promoting detection of pathogens responsible for epidemics or concerning emerging threats.

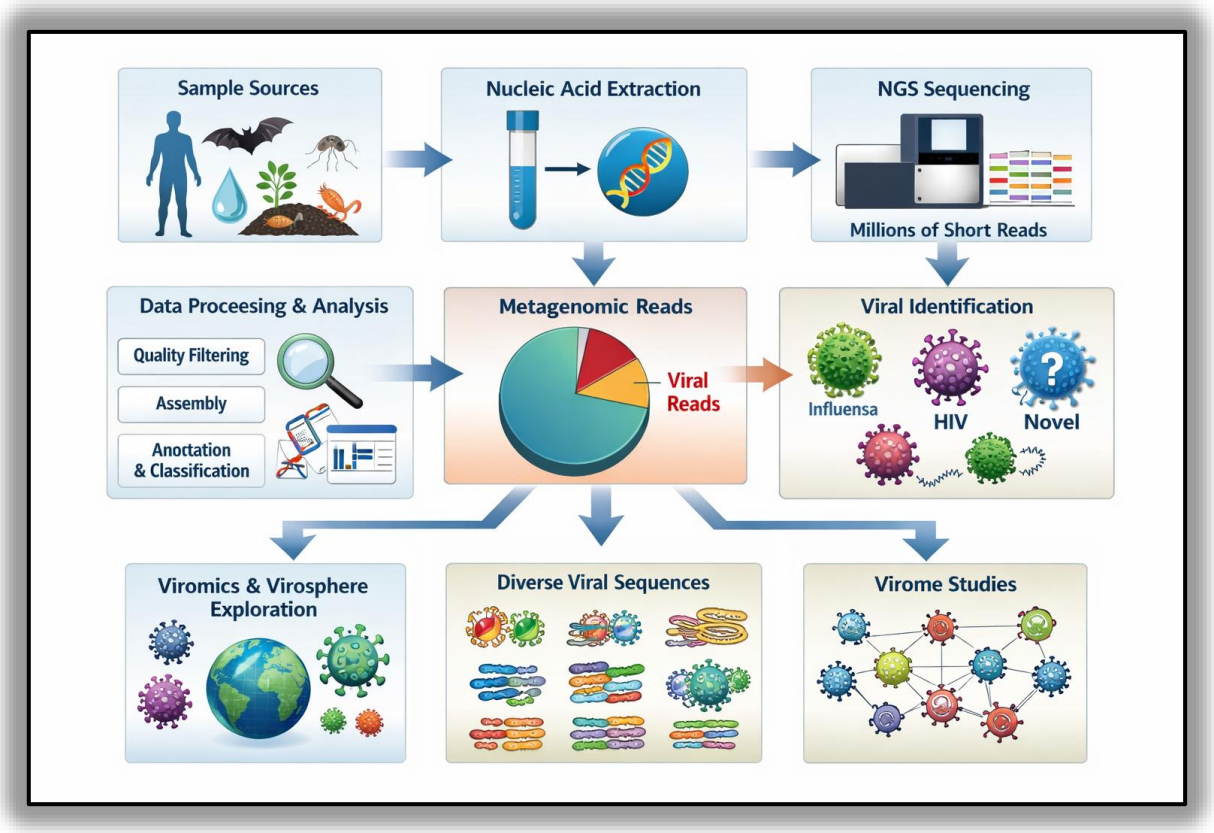

Figure1: Schematic of the Metagenomic Next-Generation Sequencing (mNGS) Workflow for Viral Detection and Virome Analysis

2.1. Analysis of Advanced Sequencing Techniques

Nucleotide sequence determination techniques continued to evolve after the first automated sequencing instruments became commercially available in the mid-1980s. Next-generation sequencing (NGS) was first used in the early 2000s; to distinguish it from earlier technologies, it is sometimes called second-generation sequencing [12]. NGS has advanced rapidly to occupy a prominent position in DNA sequencing. Seventh-generation sequencers capable of producing several terabases of data per run, [10] million reads of more than [1] kilobase in length, and [10] gigabases of data in a single hour have recently been developed [14].

NGS continues to evolve rapidly, and various systems and technologies have been developed. Table [1] lists features of the principal NGS technologies currently in widespread use and the approaches that dominate pontomic metagenomic sequencing [2]. The principal NGS systems differ greatly from one another in detection format, read length, throughput, overall workflow, and specific cost considerations; these differences profoundly influence system suitability for particular applications [12]. Moreover, different NGS platforms enable different sequencing-paradigm choices and sequence-analysis workflows. Because the technical specifications and costs of platforms continue to change, the information in Table [1] should be verified before embarking on a new project [15]

Table 1. Principal NGS Technologies, Key Features, and Dominant Sequencing Approaches

NGS Platform / Technology	Detection Format	Typical Read Length	Throughput	Dominant Sequencing Paradigms	Common Applications (Including Metagenomics)	Cost Considerations	Key Workflow Implications
Illumina	Fluorescent sequencing by synthesis	Short (50–300 bp)	Very high	Shotgun sequencing, amplicon sequencing	Genomic, transcriptomic, metagenomic profiling	Low cost per base; high instrument cost	Mature analysis pipelines; alignment-based workflows
Ion Torrent	Semiconductor-based detection	Short (100–400 bp)	Medium	Targeted sequencing, amplicon sequencing	Targeted metagenomics, microbial profiling	Moderate per-run cost	Faster runs; homopolymer error

							correction required
PacBio (SMRT)	Real-time single-molecule fluorescence	Long (10–25 kb+)	Medium	Long-read shotgun sequencing	De novo assembly, structural variation, metagenomics	High cost per base	Requires specialized long-read analysis and polishing
Oxford Nanopore	Electrical current disruption (nanopore)	Long to ultra-long (10 kb–>1 Mb)	Variable	Real-time long-read sequencing	Metagenomics, rapid pathogen detection	Flexible cost; low entry cost	Signal-level analysis; higher raw error rates
BGI / MGI	DNA nanoball sequencing with fluorescence	Short (100–150 bp)	High	Shotgun sequencing	Large-scale genomics, population metagenomics	Competitive per-base cost	Proprietary formats; Illumina-like workflows

2.2. Metagenomic Sequencing Paradigms

Metagenomic next-generation sequencing (mNGS) allows for the analysis of viral sequences extracted from the total nucleic acid pool of a clinical specimen, without pre-enrichment or amplification steps. Three main paradigms of metagenomic sequencing exist:

1) shotgun metagenomic sequencing, which facilitates the detection of multiple microorganisms, but relies on database-dependent taxonomic assignment steps;

2) targeted amplicon-based sequencing of viral genomes with primers matching highly conserved regions; and

3) hybrid approaches combining both (1) and (2) and enabling amplification-free detection of substrates belonging to a wide range of viral.

2.3. Bioinformatic Frameworks for Viral Detection

Metagenomic analysis enables pathogen detection without prior knowledge of the organism or its genome sequence. It has been applied in viral detection in clinical diagnostics, in epidemiological monitoring and environmental studies, and for monitoring vectors and reservoir species [17]. In the search for known and novel viruses in clinical samples, the metagenome-assembled genome obtained may derive from a viral genome that a predominant virus present in the sample because it is exported more efficiently to the nucleus. Furthermore, the virus must be able to generate an infectious viral particle that spreads readily from one cell to another [18].

Bioinformatics pipelines for viral detection involve four common steps: quality control (QC), read trimming and filtering, viral genome identification, and analysis of results. QC ensures high-quality data by removing adapter sequences, low-quality reads, and filtering polyclonal sequences, tailored to the specific sequencing technology [19]. Quality evaluation software tools include FastQC, Qualimap, Fastx-toolkit, and MultiQC. Viral genome identification is performed by aligning high-quality reads to the host genome using algorithms like Bowtie2, STAR, Blast, BWA, and SAMtools, followed by either alignment to reference databases such as the NCBI nt, nr, or RVDB, or de novo assembly. Aligning reads to protein databases improves sensitivity but requires more computational resources. De novo assembly reconstructs genomes from overlapping reads to produce contigs and scaffolds, a complex process demanding high coverage [2].

Proprietary software packages and open-source pipelines for analysis and interpretation of NGS data abound. Several non-commercial laboratories have developed metagenomic analysis pipelines targeted at clinical diagnostics or environmental applications, such as Taxonomer, SURPI, Metavir, MEGAN, Kraken, Virome, CensuScope, ViromeScan, Kaiju, and LMAT. Commercial software like OMICTools, CLCBio, Genedata Selector, Archetype Software, and CosmosID also facilitate bioinformatics analysis of large datasets from various HTS platforms. Analysis pipelines involve multiple steps, including quality assessment, adapter removal, trimming, reference subtraction, and sequence assembly. The choice of strategies, such as using unassembled reads or de novo assembly, affects sensitivity in detecting rare or novel viral sequences and impacts computational resource requirements [8].

3. APPLICATIONS OF MNGS IN VIRAL PATHOGEN DETECTION

Nucleic acid detection of pathogens by genomics approach is an important part of pathogen-based disease diagnosis and the development of biomarker-driven therapeutics. In recent decades, research has proven the ability of metagenomic next generation sequencing (mNGS) based detection methods to identify a wide range of potential human pathogens, including circulating adaptation and evolution of viruses of broad human concern [17]. Continuous identification of viral pathogen signatures in readily sampled environmental matrices, connecting viral reservoirs of human concern with the movement of human populations to monitor potential health threats shows that revolutionary advance in deploying this technology to the field of pathogen surveillance and public health. To such extent, it introduces the capsule state of mNGS based pathogen identification and subsequently discusses pertinent research and applications of mNGS based virome studies, pathogen identification and monitoring in the context of worldwide viral population evolution monitoring,

providing substantial evolutionary evidence on viral epidemiology globally [9]. The biosecurity and health security of the human population continues to intensify, describing the current efforts of utilizing mNGS based detection methods for human-supporting viral pathogen monitoring, emerging human in-depth virome studies and pertinent international public health efforts, exploring the future pathway and urging for action on the preservation of public rather than proprietary technology development approach in the field of pathogen detection [2]; [16].

3.1. Clinical Diagnostics and Case Studies

Recent advancements in metagenomic next-generation sequencing (mNGS) have established the technology as an innovative virology tool, providing virus detection solutions for a wide range of clinical samples, specimen types, and clinical applications. Diagnostic errors and delays frequently arise from complex, time-consuming nucleic acid extraction and detection workflows. Consequently, the identification of pathogens responsible for tropical febrile illness and other complex syndromes remains problematic, resulting in treatment delays and unfavorable patient outcomes [20]. In addition, the emergence of novel, previously unknown pathogens such as SARS-CoV-2 poses a significant public health threat.

Viral mNGS constitutes an important tool for broad-spectrum detection of viral pathogens [16]. Traditional molecular laboratory screening methods rely on the detection of known pathogens through PCR-based bidirectional or unidirectional, target-specific amplification. However, these methodologies frequently fail to detect causative pathogens in clinical samples. Comprehensive viral sequence libraries from public databases are available for mNGS broad-screening applications, enabling metabolic pathway analysis. mNGS has been applied to elucidate the causative agents of unsolved respiratory and systemic febrile diseases [21].

3.2. Surveillance of Disease Prevalence and Its Implications for Public Health

Emerging viral infections continue to threaten global health, spurring research on viral diversity and evolution. The [2014] to [2016] Ebola virus disease epidemic in West Africa, [2015] Middle East respiratory syndrome coronavirus outbreak in Korea, [2019–2020] coronavirus disease [2019] pandemic, and multi-country monkeypox resurgence highlight the need for rapid detection technologies. Early detection of emerging pathogens permits control, limiting public health risk and economic impact [8]. Viruses are significant agents of human infectious disease, yet identification remains a challenge, especially in viral meningitis and encephalitis, respiratory syndromes, febrile illness, or other clinical syndromes [16]. Traditional diagnostic testing, requiring several direct or indirect PCRs, cell culture, or serology, can take weeks. Consequently, high-throughput, comprehensive, low-cost next-generation sequencing (NGS) technologies have been implemented [2].

Metagenomic NGS enables broader pathogen detection than targeted assays, identifying not only anticipated agents but also novel or unexpected pathogens. Since [2017], this approach has been used to analyze complex clinical cases across different sample types. Viral metagenomic NGS recovered viral genomes from urine, nasal and throat swabs, and bronchoalveolar lavage fluids [22]. Detected viruses included adenoviruses, human parechoviruses, and mumps viruses, consistently known or suspected agents, for which targeted techniques were already available. Improvements in bioinformatic processing and the emergence of novel agents like avian influenza viruses and severe acute respiratory syndrome-related coronaviruses have transformed metagenomic NGS into a complementary technique for outbreak surveillance of known viruses to anticipate viral emergence and polymorphism in the human population [9].

3.3 ENVIRONMENTAL AND VECTOR-BORNE VIROLOGY

Viruses represent critical elements in both aquatic and terrestrial ecosystems, serving essential functions in the dynamics of these biomes. They regulate microbial population levels, facilitate the decomposition of organic matter, drive biogeochemical cycles, and impact nutrient transport in marine and freshwater environments. Specifically, algal viruses are significant influencers of phytoplankton populations in both lacustrine and oceanic settings, thereby affecting primary productivity and the structure of food webs [23]. In freshwater ecosystems globally, bacteriophages prevail as the predominant viral entities, modulating bacterial population densities, community composition, and their temporal variations. On the contrary, the distribution of viruses associated with sea urchins appears to be restricted across various habitats, indicating that the viral biogeography of certain eukaryotes and prokaryotes may be more influenced by localized factors than by overarching global trends [24].

Viral metagenomic next-generation sequencing (mNGS) successfully detected viral genomes from a dragonfly, the cadaver of a black-tailed rattlesnake, and urchins collected far from the laboratory, indicating that clinically important viruses could be detected from environmental sources [16]. A systemic approach was proposed to collect specimens from different environments and perform mNGS for public-health surveillance of emerging pathogens. The growing concern over vector-borne zoonotic viruses and the evolving nature of the virus classified as Group A bait have led to the establishment of a global monitoring network for uncharacterized and neglected viruses. Freshwaters host viruses relevant to vertebrates, invertebrates, algae, bacteria, and archaea. Ample sampling effort from freshwater, soil, plant, sediment, and animal specimens is needed to collect further circular DNA viruses for environmental surveillance [25].

4. METHODOLOGICAL FRAMEWORKS AND PROCESSES

Growing evidence exemplified by unprecedented outbreaks in the past two decades, such as SARS-CoV-1, MERS-CoV and Zikaviruses indicates that human pathogens derived from animal viruses represent a serious threat to public health. Until recently, viruses could be detected only through personal or environmental exposure if there was enough cause to suspect viruses as the etiology. The investigation of human or agricultural diseases caused by unknown viruses could also not be carried out without the suspected candidate viruses [26]. With the emergence of next-generation sequencing (NGS) technologies, metagenomics-based diagnostic workflows—capable of detecting previously unknown pathogens—in the framework of unbiased sequencing of nucleic acids collected from the target samples have been developed. Metagenomics sequencing for pathogen detection, which significantly simplifies conventional virus investigations, has also been included in World Health Organization (WHO) guidelines. The establishment of bioinformatic processing pipelines for read classification and consensus genome reconstruction, which is essential in high-throughput viral studies, is critical for the worldwide adoption of metagenomics [27]. Many bioinformatic pipelines based on NGS data from environmental samples, host-associated specimens and virus culture media [28] have been proposed to help tackle the scientific, health care and biosecurity issues raised by emerging viruses.

4.1. Sample Collection, Processing, and Nucleic Acid Extraction

The sample type, viral concentration achieved, and extraction methods applied have a strong influence on the quantity and quality of viral sequences recovered. Of critical importance, the composition of the sample matrix can severely inhibit downstream analyses; therefore, in complex liquids, such as serum, clarifying steps with centrifugation are almost universally recommended prior to any viral enrichment strategies [29]. Along with methods to enrich specific viral RNA subsets, ultra-clear centrifuge filters have been applied at the outset, followed by vesicle size exclusion chromatography to remove large-size contaminants within the sample. To remove any undesired contaminations, a multi-method approach for viral clarification, encompassing ultrafiltration, organic solubilization with acidic phenol, and secure-plant-based colloidal silica size-exclusion-chromatography, may also be performed on non-firewalled extraction and purification protocols in white inner-surface ldpe equipment before extraction [30]. The viral metagenomic analysis should not be restricted exclusively to the virome, as additional yet undesired viral information denoting a markedly different particle size than the targeted viruses may collectively persist within metagenomic data files and permit further entailing visualizations [28]. Notably, sample disruption and nucleic-acid extraction devices confer rapidity and efficiency to extraction methods and enhance viral-concentration yields, yet also induce significant differences upon analysis [31]. One report, investigating four commercial nucleic-acid extraction kits determined by freedom in the manipulation of RNA and solid-matrix separate exogenous ranging from double-stranded RNA to single-stranded RNA yet also DNA, underlined a strong inter-laboratory variability at the given step. Nevertheless, and essentially, many read-lengths assumptions generalize that the same reads pertains to an individual genome [32].

4.2. Strategies for Enhancing Library Development and Institutional Framework

For metagenomic NGS, sequencing libraries are constructed in a practical way to prepare specific input material, facilitate state-of-the-art detection, and fit any sequencing platform. Small fragments of DNA or RNA containing adapters need to be generated to input NGS platforms successfully [33]. For metagenomic RNA workflow, single-stranded cDNA is prepared with reverse transcription after extracting RNA from the sample [34]. Long cDNA can be directly converted to random rt-PCR cDNA without fragmentation and adapter ligation, optimizing input time. For mtDNA-like mNGS, a two-fragment approach is preferred; mtDNA is enriched and size-selected, then amplicon or short NGS libraries are constructed [31]. Prep-free is a cost-centric option that does not require library construction equipment; PCR amplification also selectively amplifies target genome after base-calling. Cambridge recommended NEBNext Ultra II or Nextera DNA Flex library construction kit. [370] µl SafBio Lyt[60] pcb Cement NGS [10]×[59] P[624184] tube-[1] is used for all Salmonella in vNPC sample [35].

4.3. Quality Control and Contamination Assessment

Quality control (QC) and contamination assessment are critical for reliable metagenomic analysis using next-generation sequencing (NGS). During library preparation, the DNA yield, size distribution, and purity of the library must be checked [36]. For an ideal NGS library preparation, DNA fragments range from [250] bp to [550] bp, and the cycle number for PCR amplification is between [6] and [15] cycles. Incorporation of indexed adapters saves sequencing time and enables variant analysis of a given sample in the same run. The quality of NGS data can be monitored through changes in the distribution of read lengths and GC contents in the sequencing process. Data can also be subjected to contamination assessment, which is important for reconstructing genomes from metagenomic data. Public genome databases are a key resource for residual reagent estimation [37]. For endogenous contamination assessment, a common approach is based on the presence and abundance of host-specific ribosomal RNAs or genomic fragments. In the case of environmental samples, open-source tools have been made available to facilitate the identification of possible

cross-sample contamination in the sequencing data, further preventing the misinterpretation of single-stranded genomic viral uncertainties from species with large genome sizes [38]. Contamination monitoring during library preparation is critical for biosecurity, especially in microbiological laboratories that handle human and animal pathogens. Contaminant DNA from reagents, surrounding environments, and even laboratory operation may be introduced into prepared samples. Next-generation sequencing (NGS) can be used to detect possible contamination in microbiology laboratories by analyzing environmental samples from various locations [39]. Sequencing of environmental nucleic acids collected from centrifuge rotors, biological test benches, and biosafety cabinets has shown that NGS is capable of monitoring contamination from different sources, including environmental microorganisms and reagent-derived nucleic acids. Incorporation of an indexed adapter into environmental-sample preparation allows monitoring of contamination throughout the entire workflow, from centrifuge to PCR-amplified libraries. Although the risk of laboratory contamination is recognized, the implementation of NGS-based assessment has proved efficient and useful [40].

4.4. Bioinformatic Pipelines and Benchmarking

The growing number of available sequencing technologies has been accompanied by an accumulation of corresponding bioinformatics pipelines for processing viral metagenomic sequencing data. In addition to many in-house pipelines developed by academic laboratories, open-source pipelines such as Taxonomer, SURPI, Metavir, MEGAN, Kraken, Virome, CensuScope, ViromeScan, Kaiju, and LMAT target different environmental or biodefense applications [41]. Companies have also embraced the opportunity, and commercial software tools (OMICTools, CLCBio, Genedata Selector, Archetype Software, CosmosID) facilitate analysis of large datasets generated by multiple sequencing platforms. All pipelines address certain general steps, including sequence read pre-processing, quality control, taxonomic or functional annotation, assembly, and linearization of sequences. The specific sequence of steps and application of optional filters vary, and important choices at different stages can significantly affect detection of rare or novel viral sequences, programme robustness, and analysis time [42]. Quality issues are generally tackled by applying tools such as FastQC and Trimmomatic, among others. After passing quality control, reads are commonly aligned to a refined set of reference sequences, although some studies instead assemble the data de novo. This choice depends on experimental design, the nature and origin of the samples, and expected read length, composition, and structure. The annotation stage is also subject to diverse criteria [43].

7. CONCLUSION

In conclusion, Next-Generation Sequencing (NGS), particularly metagenomic approaches (mNGS), has revolutionized viral discovery, overcoming limitations of traditional methods to explore the vast, underexplored virosphere. This technology enables unbiased, high-throughput detection and characterization of viruses from diverse sources, driving progress in clinical diagnostics, public health surveillance of emerging threats, and understanding viral ecology. Continuous refinement of sequencing platforms and bioinformatic frameworks is essential for maximizing the utility of mNGS in safeguarding global health.

REFERENCES

I. Koonin EV, Starokadomskyy P. Are viruses alive? The replicator paradigm sheds decisive light on an old but misguided question. Stud Hist Philos Biol Biomed Sci. 2016;59:125-134. doi:10.1016/j.shpsc.2016.02.016

II. Bassi C, Guerriero P, Pierantoni M, Callegari E, Sabbioni S. Novel Virus Identification through Metagenomics: A Systematic Review. Life (Basel). 2022;12(12):2048. doi:10.3390/life12122048

III. Conradie T, Caparros-Martin JA, Egan S, et al. Exploring the Complexity of the Human Respiratory Virome through an In Silico Analysis of Shotgun Metagenomic Data Retrieved from Public Repositories. Viruses. 2024;16(6):953. doi:10.3390/v16060953

IV. Tremblay J, Schreiber L, Greer CW. High-resolution shotgun metagenomics: the more data, the better?. Brief Bioinform. 2022;23(6):bbac443. doi:10.1093/bib/bbac443

V. Koonin EV, Kuhn JH, Dolja VV, Krupovic M. Megataxonomy and global ecology of the virosphere. ISME J. 2024;18(1):wrad042. doi:10.1093/ismejo/wrad042

VI. Angel NZ, Sullivan MJ, Alsheikh-Hussain A, et al. Metagenomics: a new frontier for routine pathology testing of gastrointestinal pathogens. Gut Pathog. 2025;17(1):4. Published 2025 Jan 18. doi:10.1186/s13099-024-00673-1

VII. Reghu R, Karunakaran RL, Balaji R, Babu A, Arumugaperumal A. Unlocking Nature's Code: The Power of Pan-Genomics in Biological Entities. MEDIN [Internet]. 2024 Dec. 24 [cited 2026 Jan. 1];. Available from: https://ojs.bonviewpress.com/index.php/MEDIN/article/view/3423

VIII. Ibañez-Lligoña M, Colomer-Castell S, González-Sánchez A, et al. Bioinformatic Tools for NGS-Based Metagenomics to Improve the Clinical Diagnosis of Emerging, Re-Emerging and New Viruses. Viruses. 2023;15(2):587. doi:10.3390/v15020587

IX. Russell T, Formiconi E, Casey M, McElroy M, Mallon PWG, Gautier VW. Viral Metagenomic

Next-Generation Sequencing for One Health Discovery and Surveillance of (Re)Emerging Viruses: A Deep Review. Int J Mol Sci. 2025;26(19):9831. doi:10.3390/ijms26199831

X. Shi T, Chen H, Liu Y, Wu Y, Lin F. Clinical applications of metagenomic next-generation sequencing in the identification of pathogens in periprosthetic joint infections: a retrospective study. J Orthop Surg Res. 2024;19(1):301. doi:10.1186/s13018-024-04745-5

XI. Kim C, Pongpanich M, Porntaveetus T. Unraveling metagenomics through long-read sequencing: a comprehensive review. J Transl Med. 2024;22(1):111. doi:10.1186/s12967-024-04917-1

XII. Kumar KR, Cowley MJ, Davis RL. Next-Generation Sequencing and Emerging Technologies. Semin Thromb Hemost. 2024;50(7):1026-1038. doi:10.1055/s-0044-1786397

XIII. Khan AS, Mallet L, Blümel J, et al. Report of the third conference on next-generation sequencing for adventitious virus detection in biologics for humans and animals. Biologicals. 2023;83:101696. doi:10.1016/j.biologicals.2023.101696

XIV. Nafea AM, Wang Y, Wang D, et al. Application of next-generation sequencing to identify different pathogens. Front Microbiol. 2024;14:1329330doi:10.3389/fmicb.2023.1329330

XV. Jiang W, Ji W, Zhang Y, et al. An Update on Detection Technologies for SARS-CoV-2 Variants of Concern. Viruses. 2022;14(11):2324. doi:10.3390/v14112324

XVI. Kufner V, Plate A, Schmutz S, et al. Two Years of Viral Metagenomics in a Tertiary Diagnostics Unit: Evaluation of the First 105 Cases. Genes (Basel). 2019;10(9):661. doi:10.3390/genes10090661

XVII. Miller S, Chiu C. The Role of Metagenomics and Next-Generation Sequencing in Infectious Disease Diagnosis. Clin Chem. 2021;68(1):115-124. doi:10.1093/clinchem/hvab173

XVIII. Lindner BG, Gerhardt K, Feistel DJ, Rodriguez-R LM, Hatt JK, Konstantinidis KT. A user's guide to the bioinformatic analysis of shotgun metagenomic sequence data for bacterial pathogen detection. Int J Food Microbiol. 2024;410:110488. doi:10.1016/j.ijfoodmicro.2023.110488

XIX. Guajardo-Leiva, S.; Chnaiderman, J.; Gaggero, A.; Díez, B. Metagenomic Insights into the Sewage RNA Virosphere of a Large City. Viruses 2020, 12, 1050. https://doi.org/10.3390/v12091050

XX. Downie DL, Rao P, David-Ferdon C, et al. Literature Review of Pathogen Agnostic Molecular Testing of Clinical Specimens From Difficult-to-Diagnose Patients: Implications for Public Health. Health Secur. 2024;22(2):93-107. doi:10.1089/hs.2023.0100

XXI. Liu H, Zhang Y, Yang J, Liu Y, Chen J. Application of mNGS in the Etiological Analysis of Lower Respiratory Tract Infections and the Prediction of Drug Resistance. Microbiol Spectr. 2022 Feb 23;10(1):e0250221. doi: 10.1128/spectrum.02502-21.

XXII. Torres Montaguth OE, Buddle S, Morfopoulou S, Breuer J. Clinical metagenomics for diagnosis and surveillance of viral pathogens. Nat Rev Microbiol. 2026 Jan;24(1):61-75. doi: 10.1038/s41579-025-01223-5.

XXIII. Wieczynski DJ, Yoshimura KM, Denison ER, Geisen S, DeBruyn JM, Shaw AJ, Weston DJ, Pelletier DA, Wilhelm SW, Gibert JP. Viral infections likely mediate microbial controls on ecosystem responses to global warming. FEMS Microbiol Ecol. 2023 Feb 28;99(3):fiad016. doi: 10.1093/femsec/fiad016. PMID: 36828391.

XXIV. Gao Y, Lu Y, Dungait JAJ, Liu J, Lin S, Jia J, Yu G. The "Regulator" Function of Viruses on Ecosystem Carbon Cycling in the Anthropocene. Front Public Health. 10.3389/fpubh.2022.858615.

XXV. Hill V, Githinji G, Vogels CBF, Bento AI, Chaguza C, Carrington CVF, Grubaugh ND. Toward a global virus genomic surveillance network. Cell Host Microbe. 2023 Jun 14;31(6):861-873. doi: 10.1016/j.chom.2023.03.003. Epub 2023 Mar 6.

XXVI. Almansour AM, Alhadlaq MA, Alzahrani KO, Mukhtar LE, Alharbi AL, Alajel SM. The silent threat: Antimicrobial-resistant pathogens in food-producing animals and their impact on public health. Microorganisms. 2023 Aug 22;11(9):2127. mdpi.com

XXVII. Parker K, Wood H, Russell JA, Yarmosh D, Shteyman A, Bagnoli J, Knight B, Aspinwall JR, Jacobs J, Werking K, Winegar R. Development and Optimization of an Unbiased, Metagenomics-Based Pathogen Detection Workflow for Infectious Disease and Biosurveillance Applications. Trop Med Infect Dis. 2023 Feb 15;8(2):121. doi: 10.3390/tropicalmed8020121.

XXVIII. Lim YW, Haynes M, Furlan M, Robertson CE, Harris JK, Rohwer F. Purifying the impure: sequencing metagenomes and metatranscriptomes from complex animal-associated samples. J Vis Exp. 2014 Dec 22;(94):52117. doi: 10.3791/52117.

XXIX. Williams ML, Olomukoro AA, Emmons RV, Godage NH, Gionfriddo E. Matrix effects demystified: Strategies for resolving challenges in analytical separations of complex samples. J Sep Sci. 2023 Dec;46(23):e2300571. doi: 10.1002/jssc.202300571.

XXX. Cabanes-Creus M, Liao SHY, Gale Navarro R, Knight M, Nazareth D, Lau NS, Ly M, Zhu E, Roca-Pinilla R, Bugallo Delgado R, Vicente AF, Baltazar

G, Westhaus A, Merjane J, Crawford M, McCaughan GW, Unzu C, González-Aseguinolaza G, Alexander IE, Pulitano C, Lisowski L. Harnessing whole human liver ex situ normothermic perfusion for preclinical AAV vector evaluation. Nat Commun. 2024 Mar 14;15(1):1876. doi: 10.1038/s41467-024-46194-y.

XXXI. Regnault B, Bigot T, Ma L, Pérot P, Temmam S, Eloit M. Deep Impact of Random Amplification and Library Construction Methods on Viral Metagenomics Results. Viruses. 2021 Feb 7;13(2):253. doi: 10.3390/v13020253.

XXXII. Zhang D, Lou X, Yan H, Pan J, Mao H, Tang H, Shu Y, Zhao Y, Liu L, Li J, Chen J, Zhang Y, Ma X. Metagenomic analysis of viral nucleic acid extraction methods in respiratory clinical samples. BMC Genomics. 2018 Oct 25;19(1):773. doi: 10.1186/s12864-018-5152-5.

XXXIII. Chen G, Qiu Y, Zhuang Q, Wang S, Wang T, Chen J, Wang K. Next-generation sequencing library preparation method for identification of RNA viruses on the Ion Torrent Sequencing Platform. Virus Genes. 2018 Aug;54(4):536-542. doi: 10.1007/s11262-018-1568-x.

XXXIV. Jia X, Hu L, Wu M, Ling Y, Wang W, Lu H, Yuan Z, Yi Z, Zhang X. A streamlined clinical metagenomic sequencing protocol for rapid pathogen identification. Sci Rep. 2021 Feb 23;11(1):4405. doi: 10.1038/s41598-021-83812-x.

XXXV. Liu Y, Ma Y. Clinical applications of metagenomics next-generation sequencing in infectious diseases. J Zhejiang Univ Sci B. 2024 May 17;25(6):471-484. doi: 10.1631/jzus.B2300029.

XXXVI. Xiao Y, Zhang L, Yang B, Li M, Ren L, Wang J. Application of next generation sequencing technology on contamination monitoring in microbiology laboratory. Biosaf Health. 2019 Jun;1(1):25-31. doi: 10.1016/j.bsheal.2019.02.003.

XXXVII. Ribarska T, Bjørnstad PM, Sundaram AYM, Gilfillan GD. Optimization of enzymatic fragmentation is crucial to maximize genome coverage: a comparison of library preparation methods for Illumina sequencing. BMC Genomics. 2022 Feb 1;23(1):92. doi: 10.1186/s12864-022-08316-y.

XXXVIII. Szabó G, Eloe-Fadrosh EA, Pett-Ridge J, Woyke T. A genomic view of Earth's Nat Rev Genet. 2026 Jan;27(1):13-31. doi: 10.1038/s41576-025-00888-1.

XXXIX. Alrasheed, Rasheed Saud, Mohammed Mahrous Al-Mahrous, Raed Mohammed Alshahrani and Nadiah Rabee Alsaadi. "Safety and Decontamination Procedures for Infectious Sample Handling." Journal Of Healthcare Sciences. 2024; 4(12): 1010-1015.

XL. Afonso CL, Afonso AM. Next-Generation Sequencing for the Detection of Microbial Agents in Avian Clinical Samples. Vet Sci. 2023 Dec 4;10(12):690. doi: 10.3390/vetsci10120690.

XLI. Lambert C, Braxton C, Charlebois RL, Deyati A, Duncan P, La Neve F, Malicki HD, Ribrioux S, Rozelle DK, Michaels B, Sun W, Yang Z, Khan AS. Considerations for Optimization of High-Throughput Sequencing Bioinformatics Pipelines for Virus Detection. Viruses. 2018 Sep 27;10(10):528. doi: 10.3390/v10100528.

XLII. Tatineni S, Hein GL. Plant Viruses of Agricultural Importance: Current and Future Perspectives of Virus Disease Management Strategies. Phytopathology. 2023 Feb;113(2):117-141. doi: 10.1094/PHYTO-05-22-0167-RVW.

XLIII. Sewe SO, Silva G, Sicat P, Seal SE, Visendi P. Trimming and Validation of Illumina Short Reads Using Trimmomatic, Trinity Assembly, and Assessment of RNA-Seq Data. Methods Mol Biol. 2022;2443:211-232. doi: 10.1007/978-1-0716-2067-0_11.